

## AMENDMENTS TO THE CLAIMS

This listing of claims will replace all prior versions, and listings, of claims in the application:

### Listing of Claims:

- 1           1. (Currently amended) A method for learning a generative model for text,  
2     comprising:  
3           receiving a current model, which contains terminal nodes representing  
4     random variables for words and can contain cluster nodes representing clusters of  
5     conceptually related words;  
6           wherein nodes in the current model are coupled together by weighted  
7     links, so that if an incoming link from a node that has fired causes a cluster node  
8     in the probabilistic model to fire with a probability proportionate to the weight of  
9     the incoming node-a cluster node in the probabilistic model fires, an outgoing-a  
10    ~~weighted~~ link from the cluster node to another node causes the other node to fire  
11    with a probability proportionate to the ~~link~~ weight of the outgoing node,  
12    otherwise, the other node does not fire;  
13           receiving a set of training documents, wherein each training document  
14    contains a set of words; and  
15           applying the set of training documents to the current model to produce a  
16    new model, wherein applying the set of training documents to the current model  
17    involves computing once for each cluster the probabilistic cost of the cluster  
18    existing in a document and triggering no words, and for each document applying  
19    this cost and subtracting the effects of words that do exist in the document.

1           2. (Original) The method of claim 1, wherein applying the set of training  
2 documents to the current model involves:  
3           applying the set of training documents to the links defined in the current  
4 model to produce functions for weights for corresponding links in the new model;  
5 and  
6           optimizing the functions to produce weights for links in the new model.

1           3. (Original) The method of claim 2, wherein for a given link, producing  
2 functions for a weight on the given link involves:  
3           producing a function for the given link for each document in the set of  
4 training documents; and  
5           multiplying the functions for each document together to produce a  
6 function to be optimized for the given link.

1           4. (Original) The method of claim 3, wherein for the given link the  
2 function for a document is an approximation of the probability of the document's  
3 terminals firing as a function of the weight on the given link, keeping all other  
4 link weights in the model constant.

1           5. (Original) The method of claim 1, wherein the method further  
2 comprises iteratively:  
3           considering the new model to be the current model; and  
4           applying training documents to the current model to produce a subsequent  
5 new model.

1           6. (Original) The method of claim 5, wherein during an initial iteration, the  
2 method further comprises generating an initial current model from a set of words  
3 by:

4           generating a universal node that is always active;  
5           generating terminal nodes representing words in the set of words; and  
6           directly linking the universal node to the terminal nodes.

1           7. (Original) The method of claim 5, wherein each iteration uses twice as  
2 many training documents as the previous iteration until all available training  
3 documents are used.

1           8. (Original) The method of claim 1, wherein producing the new model  
2 additionally involves selectively introducing new links from clusters to nodes and  
3 from clusters to clusters.

1           9. (Currently amended) The method of claim 8, wherein introducing a new  
2 link involves ~~can involve~~:  
3           considering a cluster that is assumed ~~likely~~ to be active in generating a  
4 given document;  
5           considering a new term in the given document, wherein the new term is  
6 not currently associated with the cluster; and  
7           adding the new link between the cluster and the new term.

1           10. (Currently amended) The method of claim 8, wherein introducing a  
2 new link involves ~~can involve~~:  
3           considering a first cluster that is assumed ~~likely~~ to be active in generating a  
4 given document;  
5           considering a second cluster that is assumed ~~likely~~ to be active in  
6 generating the given document, wherein the second cluster is not currently  
7 associated with the first cluster; and  
8           adding the new link between the first cluster and the second cluster.

1           11. (Original) The method of claim 1, wherein producing the new model  
2 additionally involves selectively introducing new cluster nodes into the current  
3 model.

1           12. (Original) The method of claim 11, wherein selectively introducing a  
2 new cluster node involves:  
3           examining a given document;  
4           creating the new cluster node;  
5           creating links between the new cluster node and terminals in the given  
6 document; and  
7           creating links between cluster nodes that are likely to have been involved  
8 in generating the given document and the new cluster node.

1           13. (Currently amended) The method of claim 1, wherein producing the  
2 new model involves calculating an activation for each cluster node in each  
3 document, wherein the activation for a given cluster node indicates how many  
4 links will ~~are likely to~~ fire from the given cluster node to other nodes.

1           14. (Currently amended) The method of claim 1, wherein producing the  
2 new model involves renumbering clusters in the current model to produce a  
3 cluster numbering for the new model; and  
4           wherein clusters that are ~~likely to be~~ active in generating more documents  
5 are assigned lower numbers ~~that occur earlier~~ in an identifier space, whereas  
6 clusters that are ~~likely to be~~ active in generating fewer documents are assigned  
7 higher numbers ~~that occur later~~ in the identifier space.

1           15. (Original) The method of claim 1, wherein applying a given document  
2 to the current model involves:

3 updating a summary variable for each cluster that is likely to be active in  
4 the given document, wherein the summary variable summarizes the probabilistic  
5 cost of the cluster linking to terminals not existing in the given document; and  
6 for terminals that actually do exist in the given document, canceling the  
7 effects of corresponding updates to the summary variables.

1 16 (Canceled).

1 17. (Original) The method of claim 1, wherein the probabilistic model  
2 includes a universal node that is always active and that has weighted links to  
3 terminal nodes and/or cluster nodes.

1 18. (Currently amended) A computer-readable storage medium storing  
2 instructions that when executed by a computer cause the computer to perform a  
3 method for learning a generative model for text, the method comprising:  
4 receiving a current model, which contains terminal nodes representing  
5 random variables for words and can contain cluster nodes representing clusters of  
6 conceptually related words;  
7 wherein nodes in the current model are coupled together by weighted  
8 links, so that if an incoming link from a node that has fired causes a cluster node  
9 in the probabilistic model to fire with a probability proportionate to the weight of  
10 the incoming node-a cluster node in the probabilistic model fires, an outgoing-a  
11 weighted link from the cluster node to another node causes the other node to fire  
12 with a probability proportionate to the-link weight of the outgoing node,  
13 otherwise, the other node does not fire;  
14 receiving a set of training documents, wherein each training document  
15 contains a set of words; and

16           applying the set of training documents to the current model to produce a  
17   new model, wherein applying the set of training documents to the current model  
18   involves computing once for each cluster the probabilistic cost of the cluster  
19   existing in a document and triggering no words, and for each document applying  
20   this cost and subtracting the effects of words that do exist in the document.

1           19. (Original) The computer-readable storage medium of claim 18,  
2   wherein applying the set of training documents to the current model involves:  
3           applying the set of training documents to the links defined in the current  
4   model to produce functions for weights for corresponding links in the new model;  
5   and  
6           optimizing the functions to produce weights for links in the new model.

1           20. (Original) The computer-readable storage medium of claim 19,  
2   wherein for a given link, producing functions for a weight on the given link  
3   involves:  
4           producing a function for the given link for each document in the set of  
5   training documents; and  
6           multiplying the functions for each document together to produce a  
7   function to be optimized for the given link.

1           21. (Original) The computer-readable storage medium of claim 20,  
2   wherein for the given link the function for a document is an approximation of the  
3   probability of the document's terminals firing as a function of the weight on the  
4   given link, keeping all other link weights in the model constant.

1           22. (Original) The computer-readable storage medium of claim 18,  
2   wherein the method further comprises iteratively:

3           considering the new model to be the current model; and  
4           applying training documents to the current model to produce a subsequent  
5   new model.

1           23. (Original) The computer-readable storage medium of claim 22,  
2   wherein during an initial iteration, the method further comprises generating an  
3   initial current model from a set of words by:  
4           generating a universal node that is always active;  
5           generating terminal nodes representing words in the set of words; and  
6           directly linking the universal node to the terminal nodes.

1           24. (Original) The computer-readable storage medium of claim 22,  
2   wherein each iteration uses twice as many training documents as the previous  
3   iteration until all available training documents are used.

1           25. (Original) The computer-readable storage medium of claim 18,  
2   wherein producing the new model additionally involves selectively introducing  
3   new links from clusters to nodes and from clusters to clusters.

1           26. (Original) The computer-readable storage medium of claim 25,  
2   wherein introducing a new link can involve:  
3           considering a cluster that is likely to be active in generating a given  
4   document;  
5           considering a new term in the given document, wherein the new term is  
6   not associated with the cluster; and  
7           adding the new link between the cluster and the new term.

1           27. (Original) The computer-readable storage medium of claim 25,  
2 wherein introducing a new link can involve:  
3           considering a first cluster that is likely to be active in generating a given  
4 document;  
5           considering a second cluster that is likely to be active in generating the  
6 given document, wherein the second cluster is not associated with the first cluster;  
7 and  
8           adding the new link between the first cluster and the second cluster.

1           28. (Original) The computer-readable storage medium of claim 18,  
2 wherein producing the new model additionally involves selectively introducing  
3 new cluster nodes into the current model.

1           29. (Original) The computer-readable storage medium of claim 28,  
2 wherein selectively introducing a new cluster node involves:  
3           examining a given document;  
4           creating the new cluster node;  
5           creating links between the new cluster node and terminals in the given  
6 document; and  
7           creating links between cluster nodes that are likely to have been involved  
8 in generating the given document and the new cluster node.

1           30. (Currently amended) The computer-readable storage medium of claim  
2 18, wherein producing the new model involves calculating an activation for each  
3 cluster node in each document, wherein the activation for a given cluster node  
4 indicates how many links will ~~are likely to~~ fire from the given cluster node to  
5 other nodes.



1           31. (Currently amended) The computer-readable storage medium of claim  
2 18, wherein producing the new model involves renumbering clusters in the current  
3 model to produce a cluster numbering for the new model; and  
4           wherein clusters that are ~~likely to be~~ active in generating more documents  
5 are assigned lower numbers ~~that occur earlier~~ in an identifier space, whereas  
6 clusters that are ~~likely to be~~ active in generating fewer documents are assigned  
7 higher numbers ~~that occur later~~ in the identifier space.

1           32. (Original) The computer-readable storage medium of claim 18,  
2 wherein applying a given document to the current model involves:  
3           updating a summary variable for each cluster that is likely to be active in  
4 the given document, wherein the summary variable summarizes the probabilistic  
5 cost of the cluster linking to terminals not existing in the given document; and  
6           for terminals that actually do exist in the given document, canceling the  
7 effects of corresponding updates to the summary variables.

1           33 (Canceled).

1           34. (Original) The computer-readable storage medium of claim 18,  
2 wherein the probabilistic model includes a universal node that is always active  
3 and that has weighted links to terminal nodes and/or cluster nodes.

1           35. (Currently amended) An apparatus that learns a generative model for  
2 text, comprising:  
3           a receiving mechanism configured to receive a current model, which  
4 contains terminal nodes representing random variables for words and can contain  
5 cluster nodes representing clusters of conceptually related words;

6            wherein nodes in the current model are coupled together by weighted  
7 | links, so that if an incoming link from a node that has fired causes a cluster node  
8 | in the probabilistic model to fire with a probability proportionate to the weight of  
9 | the incoming node ~~a cluster node in the probabilistic model fires, an outgoing-a~~  
10 | ~~weighted~~ link from the cluster node to another node causes the other node to fire  
11 | with a probability proportionate to the ~~link~~ weight of the outgoing node, otherwise  
12 | the other node does not fire;

13           wherein the receiving mechanism is configured to receive a set of training  
14 | documents, wherein each training document contains a set of words; and

15           a training mechanism configured to apply the set of training documents to  
16 | the current model to produce a new model, wherein applying the set of training  
17 | documents to the current model involves computing once for each cluster the  
18 | probabilistic cost of the cluster existing in a document and triggering no words,  
19 | and for each document applying this cost and subtracting the effects of words that  
20 | do exist in the document.